Lecture Notes 1: Matrix Algebra
Part A: Vectors and Matrices

Peter J. Hammond
email: hammond@stanford.edu

Autumn 2012, revised 2014

Lecture Outline

Solving Two Equations in Two Unknowns

Vectors

Matrices

Example of Two Equations in Two Unknowns

It is easy to check that

$$\left.\begin{aligned} x + y &= 10 \\ x - y &= 6 \end{aligned}\right\} \Longrightarrow x = 8,\ y = 2$$

More generally, one can:

1. add the two equations, to eliminate y;
2. subtract the second equation from the first, to eliminate x.

This leads to the following transformation

$$\left.\begin{aligned} x + y &= b_1 \\ x - y &= b_2 \end{aligned}\right\} \Longrightarrow \left\{\begin{aligned} 2x &= b_1 + b_2 \\ 2y &= b_1 - b_2 \end{aligned}\right.$$

of the two equation system with general right-hand sides.

Obviously the solution is

$$x = \tfrac{1}{2}(b_1 + b_2),\ y = \tfrac{1}{2}(b_1 - b_2)$$

Using Matrix Notation, I

Matrix notation allows the two equations

$$1x + 1y = b_1$$
$$1x - 1y = b_2$$

to be expressed as

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

or as $\mathbf{Az} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Here $\mathbf{A}, \mathbf{z}, \mathbf{b}$ are respectively: (i) the coefficient matrix; (ii) the vector of unknowns; (iii) the vector of right-hand sides.

Using Matrix Notation, II

Also, the solution $x = \frac{1}{2}(b_1 + b_2),\ y = \frac{1}{2}(b_1 - b_2)$ can be expressed as

$$x = \tfrac{1}{2}b_1 + \tfrac{1}{2}b_2$$
$$y = \tfrac{1}{2}b_1 - \tfrac{1}{2}b_2$$

or as

$$\mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathbf{Cb}, \quad \text{where} \quad \mathbf{C} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Two General Equations

Consider the general system

$$\begin{array}{rcccl} ax + by & = & u & = & 1u + 0v \\ cx + dy & = & v & = & 0u + 1v \end{array}$$

of two equations in two unknowns, filled in with 1s and 0s.

In matrix form, these equations can be written as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

In case $a \neq 0$, we can eliminate x from the second equation by adding $-c/a$ times the first row to the second.

After defining the scalar constant $D := a[d + (-c/a)b] = ad - bc$, then clearing fractions, we obtain the new equality

$$\begin{pmatrix} a & b \\ 0 & D/a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -c/a & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

Two General Equations, Subcase 1A

In **Subcase 1A** when $D := ad - bc \neq 0$,
multiply the second row by a to obtain

$$\begin{pmatrix} a & b \\ 0 & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -c & a \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

Adding $-b/D$ times the second row to the first yields

$$\begin{pmatrix} a & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 + (bc/D) & -ab/D \\ -c & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

Recognizing that $1 + (bc/D) = (D + bc)/D = ad/D$,
then dividing the two rows/equations by a and D respectively,
we obtain

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{D} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

which implies the unique solution

$$x = (1/D)(du - bv) \quad \text{and} \quad y = (1/D)(av - cu)$$

Two General Equations, Subcase 1B

In **Subcase 1B** when $D := ad - bc = 0$,
the multiplier $-ab/D$ is undefined and the system

$$\begin{pmatrix} a & b \\ 0 & D/a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -c/a & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

collapses to

$$\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v - c/a \end{pmatrix}.$$

This leaves us with two "subsubcases":

if $c \neq av$, then the left-hand side of the second equation is 0,
but the right-hand side is non-zero,
so there is no solution;

if $c = av$, then the second equation reduces to $0 = 0$,
and there is a continuum of solutions
satisfying the one remaining equation $ax + by = u$,
or $x = (u - by)/a$ where y is any real number.

Two General Equations, Case 2

In the final case when $a = 0$, simply interchanging the two equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

gives

$$\begin{pmatrix} c & d \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix}.$$

Provided that $b \neq 0$, one has $y = u/b$ and,
assuming that $c \neq 0$, also $x = (v - dy)/c = (bv - du)/bc$.

On the other hand, if $b = 0$,
we are back with two possibilities like those of Subcase 1B.

Outline

Vectors and Inner Products

Let $\mathbf{x} = (x_i)_{i=1}^m \in \mathbb{R}^m$ denote a column m-vector of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Its transpose is the row m-vector

$$\mathbf{x}^\top = (x_1, x_2, \dots, x_m).$$

Given a column m-vector $\mathbf{x}$ and row n-vector $\mathbf{y}^\top = (y_j)_{j=1}^n \in \mathbb{R}^n$ where $m = n$, the dot or scalar or inner product is defined as

$$\mathbf{y}^\top \mathbf{x} := \mathbf{y} \cdot \mathbf{x} := \sum\nolimits_{i=1}^n y_i x_i.$$

But when $m \neq n$, the scalar product is not defined.

Exercise on Quadratic Forms

Exercise

Consider the **quadratic form** $f(\mathbf{w}) := \mathbf{w}^\top \mathbf{w}$
as a function $f : \mathbb{R}^n \to \mathbb{R}$ *of the column* n*-vector* $\mathbf{w}$.

Explain why $f(\mathbf{w}) \geq 0$ *for all* $\mathbf{w} \in \mathbb{R}^n$,
with equality if and only if $\mathbf{w} = \mathbf{0}$,
where $\mathbf{0}$ *denotes the zero vector of* $\mathbb{R}^n$.

Net Quantity Vectors

Suppose there are n commodities numbered from $i = 1$ to n.

Each component q_i of the net quantity vector $\mathbf{q} = (q_i)_{i=1}^n \in \mathbb{R}^n$ represents the quantity of the ith commodity.

Often each such quantity is non-negative.

But general equilibrium theory often uses only by the sign of q_i to distinguish between

- a consumer's demands and supplies of the ith commodity;
- or a producer's outputs and inputs of the ith commodity.

This sign is taken to be

positive for demands or outputs;

negative for supplies or inputs.

In fact, q_i is taken to be

- the consumer's net demand for the ith commodity;
- the producer's net supply or net outputs of the ith commodity.

Then $\mathbf{q}$ is the net quantity vector.

Price Vectors

Each component p_i of the (row) price vector $\mathbf{p}^\top \in \mathbb{R}^n$ indicates the price per unit of commodity i.

Then the scalar product

$$\mathbf{p}^\top \mathbf{q} = \mathbf{p} \cdot \mathbf{q} = \sum_{i=1}^{n} p_i q_i$$

is the total value of the net quantity vector $\mathbf{q}$ evaluated at the price vector $\mathbf{p}$.

In particular, $\mathbf{p}^\top \mathbf{q}$ indicates

- the net profit (or minus the net loss) for a producer;
- the net dissaving for a consumer.

Outline

Definitions

Consider any two n-vectors $\mathbf{x} = (x_i)_{i=1}^n$ and $\mathbf{y} = (y_i)_{i=1}^n$ in $\mathbb{R}^n$.

Their sum $\mathbf{s} := \mathbf{x} + \mathbf{y}$ and difference $\mathbf{d} := \mathbf{x} - \mathbf{y}$
are constructed by adding or subtracting the vectors
component by component — i.e.,

$$\mathbf{s} = (x_i + y_i)_{i=1}^n \quad \text{and} \quad \mathbf{d} = (x_i - y_i)_{i=1}^n$$

The scalar product $\lambda \mathbf{x}$ of any scalar $\lambda \in \mathbb{R}$
and vector $\mathbf{x} = (x_i)_{i=1}^n \in \mathbb{R}^n$ is constructed by multiplying
each component of the vector $\mathbf{x}$ by the scalar λ — i.e.,

$$\lambda \mathbf{x} = (\lambda x_i)_{i=1}^n$$

Algebraic Fields

Definition

An algebraic field $(\mathbb{F}, +, \cdot)$ of scalars is a set $\mathbb{F}$ that, together with the two binary operations $+$ of addition and $\cdot$ of multiplication, satisfies the following axioms for all $a, b, c \in \mathbb{F}$:

1. $\mathbb{F}$ is closed under $+$ and $\cdot$: both $a + b$ and $a \cdot b$ are in $\mathbb{F}$.
2. $+$ and $\cdot$ are associative: both $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
3. $+$ and $\cdot$ both commute: $a + b = b + a$ and $a \cdot b = b \cdot a$.
4. There are identity elements $0, 1 \in \mathbb{F}$ for $+$ and $\cdot$ respectively: for all $a \in \mathbb{F}$, one has $a + 0 = a$ and $1 \cdot a = a$, with $0 \neq 1$.
5. There are inverse operations $-$ for $+$ and $^{-1}$ for $\cdot$ such that: $a + (-a) = 0$ and $a \cdot a^{-1} = 1$ provided that $a \neq 0$.
6. The distributive law: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

Examples of Algebraic Fields

Exercise

Verify that the following well known sets are algebraic fields:

- $\mathbb{R}$, *the set of all real numbers;*
- $\mathbb{Q}$, *the set of all rational numbers*
  — i.e., those that can be expressed as the ratio $r = p/q$
  of integers $p, q \in \mathbb{Z}$ *with* $q \neq 0$.
  (Check that $\mathbb{Q}$ *is closed under addition and multiplication,*
  and that each non-zero rational
  has a rational multiplicative inverse.)
- $\mathbb{C}$, *the set of all complex numbers*
  — i.e., those that can be expressed as $c = a + ib$,
  where $a, b \in \mathbb{R}$ *and* i *is defined to satisfy* $i^2 = -1$.
- *the set of all rational complex numbers*
  — i.e., those that can be expressed as $c = a + ib$,
  where $a, b \in \mathbb{Q}$ *and* i *is defined to satisfy* $i^2 = -1$.

General Vector Spaces

Definition

A vector space V over an algebraic field $\mathbb{F}$
is a combination $\langle V, \mathbb{F}, +, \cdot \rangle$ of:

- a set V of vectors;
- the field $\mathbb{F}$ of scalars;
- the binary operation
$$V \times V \ni (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v} \in V$$
  of vector addition
- the binary operation
$$\mathbb{F} \times V \ni (\alpha, \mathbf{u}) \mapsto \alpha \mathbf{u} \in V$$
  of scalar multiplication

which are required to satisfy
all of the following eight vector space axioms.

Eight Vector Space Axioms

1. Addition is associative: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
2. Addition is commutative: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. Additive identity: There exists a zero vector $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.
4. Additive inverse: For every $\mathbf{v} \in V$, there exists an additive inverse $-\mathbf{v} \in V$ of $\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.
5. Scalar multiplication is distributive w.r.t. vector addition: $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$
6. Scalar multiplication is distributive w.r.t. field addition: $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$
7. Scalar and field multiplication are compatible: $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$
8. $1 \in \mathbb{F}$ is an identity element for scalar multiplication: $1\mathbf{v} = \mathbf{v}$

Some Finite Dimensional Vector Spaces

Exercise

Given an arbitrary algebraic field $\mathbb{F}$*, let* $\mathbb{F}^n$ *denote the space of all lists* $\langle a_i \rangle_{i=1}^n$ *of* n *elements* $a_i \in \mathbb{F}$ *— i.e., the* n*-fold Cartesian product of* $\mathbb{F}$ *with itself.*
Show how to construct the respective binary operations

$$\mathbb{F}^n \times \mathbb{F}^n \ni (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathbf{y} \in \mathbb{F}^n$$
$$\mathbb{F} \times \mathbb{F}^n \ni (\lambda, \mathbf{x}) \mapsto \lambda \mathbf{x} \in \mathbb{F}^n$$

of addition and scalar multiplication so that $(\mathbb{F}^n, \mathbb{F}, +, \times)$ *is a vector space.*

Show too that subtraction and division by a (non-zero) scalar can be defined by $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1)\mathbf{w}$ *and* $\mathbf{v}/\alpha = (1/\alpha)\mathbf{v}$*.*

From now on we consider real vector spaces over the real field $\mathbb{R}$, and especially the space $(\mathbb{R}^n, \mathbb{R}, +, \times)$ of n-vectors over $\mathbb{R}$.

Outline

Linear Functions: Definition

Definition

A linear combination of vectors is the weighted sum $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h$, where $\mathbf{x}^h \in V$ and $\lambda_h \in \mathbb{F}$ for $h = 1, 2, \ldots, k$.

Exercise

By induction on k, show that the vector space axioms imply that any linear combination of vectors in V must also belong to V.

Definition

A function $V \ni \mathbf{u} \mapsto f(\mathbf{u}) \in \mathbb{F}$ is linear provided that

$$f(\lambda \mathbf{u} + \mu \mathbf{v}) = \lambda f(\mathbf{u}) + \mu f(\mathbf{v})$$

whenever $\mathbf{u}, \mathbf{v} \in V$ and $\lambda, \mu \in V$.

Key Properties of Linear Functions

Exercise

Use induction on k to show that if the function $f : V \to \mathbb{F}$ is linear, then

$$f\left(\sum_{h=1}^{k} \lambda_h \mathbf{x}^h\right) = \sum_{h=1}^{k} \lambda_h f(\mathbf{x}^h)$$

for all linear combinations $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h$ in V — i.e., f preserves linear combinations.

Exercise

In case $V = \mathbb{R}^n$ and $\mathbb{F} = \mathbb{R}$, show that any linear function is homogeneous of degree 1, meaning that $f(\lambda \mathbf{v}) = \lambda f(\mathbf{v})$ for all $\lambda \in \mathbb{R}$ and all $\mathbf{v} \in \mathbb{R}^n$.

What is the corresponding property in case $V = \mathbb{Q}^n$ and $\mathbb{F} = \mathbb{Q}$?

Affine Functions

Definition

A function $g : V \to \mathbb{F}$ is said to be affine
if there is a scalar additive constant $\alpha \in \mathbb{F}$
and a linear function $f : V \to \mathbb{F}$ such that $g(\mathbf{v}) \equiv \alpha + f(\mathbf{v})$.

Exercise

Under what conditions is an affine function $g : \mathbb{R} \to \mathbb{R}$ linear when its domain $\mathbb{R}$ is regarded as a vector space?

An Economic Aggregation Theorem

Suppose that a finite population of households $h \in H$
with respective non-negative incomes $y_h \in \mathbb{Q}_+$ $(h \in H)$
have non-negative demands $x_h \in \mathbb{R}$ $(h \in H)$
which depend on household income via a function $y_h \mapsto f_h(y_h)$.

Given total income $Y := \sum_h y_h$, under what conditions
can their total demand $X := \sum_h x_h = \sum_h f_h(y_h)$
be expressed as a function $X = F(Y)$ of Y alone?

The answer is an implication of Cauchy's functional equation.

In this context the theorem asserts that this aggregation condition
implies that the functions f_h $(h \in H)$ and F must be co-affine.

This means there exists a common multiplicative constant $\rho \in \mathbb{R}$,
along with additive constants α_h $(h \in H)$ and A, such that

$$f_h(y_h) \equiv \alpha_h + \rho y_h \ (h \in H) \text{ and } F(Y) \equiv A + \rho Y$$

Cauchy's Functional Equation: Proof of Sufficiency

Theorem

Except in the trivial case when H has only one member, Cauchy's functional equation $F(\sum_h y_h) \equiv \sum_h f_h(y_h)$ is satisfied for functions $F, f_h : \mathbb{Q} \to \mathbb{R}$ if and only if:

1. *there exists a single function $\phi : \mathbb{Q} \to \mathbb{R}$ such that*

   $$F(q) = F(0) + \phi(q) \text{ and } f_h(q) = f_h(0) + \phi(q) \text{ for all } h \in H$$

2. *the function $\phi : \mathbb{Q} \to \mathbb{R}$ is linear, implying that the functions F and f_h are co-affine.*

Proof.

Suppose $f_h(y_h) \equiv \alpha_h + \rho y_h$ for all $h \in H$, and $F(Y) \equiv A + \rho Y$. Then Cauchy's functional equation $F(\sum_h y_h) \equiv \sum_h f_h(y_h)$ is obviously satisfied provided that $A = \sum_h \alpha_h$. □

Cauchy's Functional Equation: Beginning the Proof

Lemma

The mapping $\mathbb{Q} \ni q \mapsto \phi(q) := F(q) - F(0) \in \mathbb{R}$ *must satisfy;*

1. $\phi(q) \equiv f_i(q) - f_i(0)$ *for all* $i \in H$ *and* $q \in \mathbb{Q}$*;*
2. $\phi(q + q') \equiv \phi(q) + \phi(q')$ *for all* $q, q' \in \mathbb{Q}$*.*

Proof.

To prove part 1, consider any $i \in H$ and all $q \in \mathbb{Q}$.

Note that Cauchy's equation $F(\sum_h y_h) \equiv \sum_h f_h(y_h)$
implies that $F(q) = f_i(q) + \sum_{h \neq i} f_h(0)$
and also $F(0) = f_i(0) + \sum_{h \neq i} f_h(0)$.

Now subtract the second equation from the first to obtain

$$\phi(q) = F(q) - F(0) = f_i(q) - f_i(0)$$

□

Cauchy's Functional Equation: Continuing the Proof

Lemma

The mapping $\mathbb{Q} \ni q \mapsto \phi(q) := F(q) - F(0) \in \mathbb{R}$ *must satisfy;*

1. $\phi(q) \equiv f_i(q) - f_i(0)$ *for all* $i \in H$ *and* $q \in \mathbb{Q}$*;*
2. $\phi(q + q') \equiv \phi(q) + \phi(q')$ *for all* $q, q' \in \mathbb{Q}$*.*

Proof.

To prove part 2, consider any $i, j \in H$ and any $q, q' \in \mathbb{Q}$.

Note that Cauchy's equation $F(\sum_h y_h) \equiv \sum_h f_h(y_h)$ implies that

$$\begin{aligned} F(q + q') &= f_i(q) + f_j(q') + \textstyle\sum_{h \in H \setminus \{i,j\}} f_h(0) \\ F(0) &= f_i(0) + f_j(0) + \textstyle\sum_{h \in H \setminus \{i,j\}} f_h(0) \end{aligned}$$

Now subtract the second equation from the first,
and use the equation $\phi(q) = F(q) - F(0) = f_i(q) - f_i(0)$,
to obtain $\phi(q + q') = \phi(q) + \phi(q')$.

□

Cauchy's Functional Equation: Resuming the Proof

Because $\phi(q + q') \equiv \phi(q) + \phi(q')$,
for any $k \in \mathbb{N}$ one has $\phi(kq) = \phi((k-1)q) + \phi(q)$.

As an induction hypothesis, which is trivially true for $k = 2$, suppose that $\phi((k-1)q) = (k-1)\phi(q)$.

Confirming the induction step, the hypothesis implies that

$$\phi(kq) = \phi((k-1)q) + \phi(q) = (k-1)\phi(q) + \phi(q) = k\phi(q)$$

So $\phi(kq) = k\phi(q)$ for every $k \in \mathbb{N}$ and every $q \in \mathbb{Q}$.

Putting $q' = kq$ implies that $\phi(q') = k\phi(q'/k)$.

Interchanging q and q', it follows that $\phi(q/k) = (1/k)\phi(q)$.

Cauchy's Functional Equation: Completing the Proof

So far we have proved that, for every $k \in \mathbb{N}$ and every $q \in \mathbb{Q}$, one has both $\phi(kq) = k\phi(q)$ and $\phi(q/k) = (1/k)\phi(q)$.

Hence, for every rational $r = m/n \in \mathbb{Q}$
one has $\phi(mq/n) = m\phi(q/n) = (m/n)\phi(q)$ and so $\phi(rq) = r\phi(q)$.

In particular, $\phi(r) = r\phi(1)$, so ϕ is linear on its domain $\mathbb{Q}$ (though not on the whole of $\mathbb{R}$ without additional assumptions such as continuity or monotonicity).

The rest of the proof is routine checking of definitions.

Outline

Norm as Length

Pythagoras's theorem implies that the length of the typical vector $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ is $\sqrt{x_1^2 + x_2^2}$ or, perhaps less clumsily, $(x_1^2 + x_2^2)^{1/2}$.

In $\mathbb{R}^3$, the same result implies that the length of the typical vector $\mathbf{x} = (x_1, x_2, x_3)$ is

$$\left[\left((x_1^2 + x_2^2)^{1/2}\right)^2 + x_3^2\right]^{1/2} = (x_1^2 + x_2^2 + x_3^2)^{1/2}.$$

An obvious extension to $\mathbb{R}^n$ is the following:

Definition

The length of the typical n-vector $\mathbf{x} = (x_i)_{i=1}^n \in \mathbb{R}^n$ is its (Euclidean) norm

$$\|\mathbf{x}\| := \left(\sum_{i=1}^n x_i^2\right)^{1/2} = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

Unit n-Vectors, the Unit Sphere, and Unit Ball

Definition

A unit vector $\mathbf{u} \in \mathbb{R}^n$ is a vector with unit norm
— i.e., its components satisfy $\sum_{i=1}^n u_i^2 = \|\mathbf{u}\| = 1$.

The set of all such unit vectors forms a surface called the unit sphere of dimension $n-1$ (one less than n because of the defining equation), defined as

$$S^{n-1} := \{\mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1\}$$

The unit ball $B \subset \mathbb{R}^n$ is the solid set

$$B := \{\mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$$

of all points bounded by the surface of the unit sphere $S^{n-1} \subset \mathbb{R}^n$.

Cauchy–Schwartz Inequality

Theorem

For all pairs $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, one has $|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \|\mathbf{b}\|$.

Proof.

Define the function $\mathbb{R} \ni \xi \mapsto f(\xi) := \sum_{i=1}^n (a_i \xi + b_i)^2 \in \mathbb{R}$.

Clearly f is the quadratic form $f(\xi) \equiv A\xi^2 + B\xi + C$
where $A := \sum_{i=1}^n a_i^2 = \|\mathbf{a}\|^2$, $B := 2 \sum_{i=1}^n a_i b_i = 2\mathbf{a} \cdot \mathbf{b}$,
and $C := \sum_{i=1}^n b_i^2 = \|\mathbf{b}\|^2$.

There is a trivial case when $A = 0$ because $\mathbf{a} = \mathbf{0}$.

Otherwise, $A > 0$ and so completing the square gives

$$f(\xi) \equiv A\xi^2 + B\xi + C = A[\xi + (B/2A)]^2 + C - B^2/4A$$

But the definition of f implies that $f(\xi) \geq 0$ for all $\xi \in \mathbb{R}$,
so $0 \leq f(-B/2A) = C - B^2/4A$, implying that $\frac{1}{4}B^2 \leq AC$
and so $|\mathbf{a} \cdot \mathbf{b}| = |\frac{1}{2}B| \leq \sqrt{AC} = \|\mathbf{a}\| \|\mathbf{b}\|$. □

Outline

The Angle Between Two Vectors

Consider the triangle in $\mathbb{R}^n$ whose vertices are the vectors $\mathbf{x}, \mathbf{y}, \mathbf{0}$.

Its three sides or edges have respective lengths $\|\mathbf{x}\|$, $\|\mathbf{y}\|$, $\|\mathbf{x} - \mathbf{y}\|$, where the last follows from the parallelogram law.

Note that $\|\mathbf{x} - \mathbf{y}\|^2 \lesseqgtr \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ according as the angle at $\mathbf{0}$ is: (i) acute; (ii) a right angle; (iii) obtuse. But

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 &= \sum_{i=1}^n (x_i - y_i)^2 - \sum_{i=1}^n x_i^2 - \sum_{i=1}^n y_i^2 \\ &= -2 \sum_{i=1}^n x_i y_i = -2\mathbf{x} \cdot \mathbf{y} \end{aligned}$$

So the three cases (i)–(iii) occur according as $\mathbf{x} \cdot \mathbf{y} \gtreqless 0$.

Using the Cauchy–Schwartz inequality, one can define the angle between $\mathbf{x}$ and $\mathbf{y}$ as the unique solution $\theta = \arccos(\mathbf{x} \cdot \mathbf{y} / \|\mathbf{x}\| \|\mathbf{y}\|)$ in the interval $[0, \pi]$ of the equation $\cos\theta = \mathbf{x} \cdot \mathbf{y} / \|\mathbf{x}\| \|\mathbf{y}\| \in [-1, 1]$.

Orthogonal and Orthonormal Sets of Vectors

Case (ii) suggests defining two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ as orthogonal iff $\mathbf{x} \cdot \mathbf{y} = 0$.

A set of k vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is said to be:

- pairwise orthogonal just in case $\mathbf{x} \cdot \mathbf{y} = 0$ whenever $j \neq i$;
- orthonormal just in case, in addition, all k elements of the set are unit vectors.

On the set $\{1, 2, \ldots, n\}$, define the Kronecker delta function

$$\{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\} \ni (i, j) \mapsto \delta_{ij} \in \{0, 1\}$$

by

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Then the set of k vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is orthonormal if and only if $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$ for all pairs $i, j \in \{1, 2, \ldots, k\}$.

Outline

The Canonical Basis of $\mathbb{R}^n$

Example

A prominent orthonormal set is the canonical basis of $\mathbb{R}^n$, defined as the set of n different n-vectors $\mathbf{e}^i$ $(i = 1, 2, \ldots, n)$ whose respective components $(e^i_j)^n_{j=1}$ satisfy $e^i_j = \delta_{ij}$ for all $j \in \{1, 2, \ldots, n\}$.

Exercise

Show that each n-vector $\mathbf{x} = (x_i)^n_{i=1}$ *is a linear combination*

$$\mathbf{x} = (x_i)^n_{i=1} = \sum\nolimits_{i=1}^{n} x_i \mathbf{e}_i$$

of the canonical basis vectors, with the multiplier attached to each basis vector $\mathbf{e}_i$ *equal to the respective component* x_i $(i = 1, 2, \ldots, n)$*.*

The Canonical Basis in Commodity Space

Example

Consider the case when each vector $\mathbf{x} \in \mathbb{R}^n$ is a quantity vector, whose components are $(x_i)_{i=1}^n$, where x_i indicates the net quantity of commodity i.

Then the ith unit vector $\mathbf{e}^i$ of the canonical basis of $\mathbb{R}^n$ represents a commodity bundle that consists of one unit of commodity i, but nothing of every other commodity.

In case the row vector $\mathbf{p}^\top \in \mathbb{R}^n$ is a price vector for the same list of n commodities, the value $\mathbf{p}^\top \mathbf{e}^i$ of the ith unit vector $\mathbf{e}^i$ must equal p_i, the price (of one unit) of the ith commodity.

Linear Functions

Theorem

If the function $f : \mathbb{R}^n \to \mathbb{R}$ *is linear,*
there exists $\mathbf{y} \in \mathbb{R}^n$ *such that* $f(\mathbf{x}) = \mathbf{y}^\top \mathbf{x}$.

Proof.

Because $\mathbf{x}$ equals the linear combination $\sum_{i=1}^n x_i \mathbf{e}_i$
of the n canonical basis vectors, linearity of f implies that

$$f(\mathbf{x}) = f\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) = \sum_{i=1}^n x_i f(\mathbf{e}_i) = \mathbf{y}^\top \mathbf{x}$$

where $\mathbf{y}$ is the column vector whose components
are $y_i = f(\mathbf{e}_i)$ for $i = 1, 2, \ldots, n$. □

Linear Transformations: Definition

Definition

The vector-valued function

$$\mathbb{R}^n \ni \mathbf{x} \mapsto F(\mathbf{x}) = (F_i(\mathbf{x}))_{i=1}^m \in \mathbb{R}^m$$

is a linear transformation
just in case each component function $F_i : \mathbb{R}^n \to \mathbb{R}$ is linear
— or equivalently, iff $F(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda F(\mathbf{x}) + \mu F(\mathbf{y})$
for every linear combination $\lambda \mathbf{x} + \mu \mathbf{y}$ of every pair $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Linear Transformations: Representation

Theorem

For any linear transformation $F : \mathbb{R}^n \to \mathbb{R}^m$,
there exist vectors $\mathbf{y}_i \in \mathbb{R}^m$ *for* $i = 1, 2, \ldots, n$
such that each component function satisfies $F_i(\mathbf{x}) = \mathbf{y}_i^\top \mathbf{x}$.

Proof.

Because $\mathbf{x}$ equals the linear combination $\sum_{i=1}^n x_i \mathbf{e}_i$
of the n canonical basis vectors, linearity of F_i implies that

$$F_i(\mathbf{x}) = F_i\left(\sum_{j=1}^n x_j \mathbf{e}_j\right) = \sum_{j=1}^n x_j F_i(\mathbf{e}_j) = \mathbf{y}_i^\top \mathbf{x}$$

where $\mathbf{y}_i^\top$ is the row vector whose components
are $(\mathbf{y}_i)_j = F_i(\mathbf{e}_j)$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. □

Consider the $m \times n$ array whose n columns
are the m-vectors $F(\mathbf{e}_j) = (F_i(\mathbf{e}_j))_{i=1}^m$ for $j = 1, 2, \ldots, n$.
This is a matrix representation of the linear transformation F.

Outline

Linear Combinations and Dependence: Definitions

Definition

A linear combination of the finite set $\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$ of vectors is the scalar weighted sum $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h$,
where $\lambda_h \in \mathbb{R}$ for $h = 1, 2, \ldots, k$.

The finite set $\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$ of vectors is linearly independent just in case the only solution of the equation $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h = 0$ is the trivial solution $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$.

Alternatively, if the equation has a non-trivial solution, then the set of vectors is linearly dependent.

Characterizing Linear Dependence

Theorem

The finite set $\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$ of vectors is linearly dependent if and only if at least one of the vectors, say $\mathbf{x}^1$ after reordering, can be expressed as a linear combination of the others — i.e., there exist scalars α^h ($h = 2, 3, \ldots, k$) such that $\mathbf{x}^1 = \sum_{h=2}^{k} \alpha_h \mathbf{x}^h$.

Proof.

If $\mathbf{x}^1 = \sum_{h=2}^{k} \alpha_h \mathbf{x}^h$, then $(-1)\mathbf{x}^1 + \sum_{h=2}^{k} \alpha_h \mathbf{x}^h = 0$, so $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h = 0$ has a non-trivial solution.

Conversely, suppose $\sum_{h=1}^{k} \lambda_h \mathbf{x}^h = 0$ has a non-trivial solution.

After reordering, we can suppose that $\lambda_1 \neq 0$.

Then $\mathbf{x}^1 = \sum_{h=2}^{k} \alpha_h \mathbf{x}^h$,
where $\alpha_h = -\lambda_h / \lambda_1$ for $h = 2, 3, \ldots, k$. □

Dimension

Definition

The dimension of a vector space V is the size of any maximal set of linearly independent vectors, if this number is finite.

Otherwise, if there is an infinite set of linearly independent vectors, the dimension is infinite.

Exercise

Show that the canonical basis of $\mathbb{R}^n$ is linearly independent.

Example

The previous exercise shows that the dimension of $\mathbb{R}^n$ is at least n.

Later, we will show that any set of $k > n$ vectors in $\mathbb{R}^n$ is linearly dependent.

This implies that the dimension of $\mathbb{R}^n$ is exactly n.

Outline

Matrices as Rectangular Arrays

An $m \times n$ matrix $\mathbf{A} = (a_{ij})^{m \times n}$ is a (rectangular) array

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = ((a_{ij})_{i=1}^m)_{j=1}^n = ((a_{ij})_{j=1}^n)_{i=1}^m$$

An $m \times 1$ matrix is a column vector with m rows and 1 column.

A $1 \times n$ matrix is a row vector with 1 row and n columns.

The $m \times n$ matrix $\mathbf{A}$ consists of:

n columns in the form of m-vectors
$\mathbf{a}_j = (a_{ij})_{i=1}^m \in \mathbb{R}^m$ for $j = 1, 2, \dots, n$;

m rows in the form of n-vectors
$\mathbf{a}_i^\top = (a_{ij})_{j=1}^n \in \mathbb{R}^n$ for $i = 1, 2, \dots, m$
which are transposes of column vectors.

The Transpose of a Matrix

The transpose of the $m \times n$ matrix $\mathbf{A} = (a_{ij})_{m \times n}$
is the $n \times m$ matrix

$$\mathbf{A}^\top = (a_{ij}^\top)_{n \times m} = (a_{ji})_{n \times m} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}$$

which results from transforming
each column m-vector $\mathbf{a}_j = (a_{ij})_{i=1}^m$ $(j = 1, 2, \dots, n)$ of $\mathbf{A}$
into the corresponding row m-vector $\mathbf{a}_j^\top = (a_{ji}^\top)_{i=1}^m$ of $\mathbf{A}^\top$.

Equivalently, for each $i = 1, 2, \dots, m$,
the ith row n-vector $\mathbf{a}_i^\top = (a_{ij})_{j=1}^n$ of $\mathbf{A}$
is transformed into the ith column n-vector $\mathbf{a}_i = (a_{ji})_{j=1}^n$ of $\mathbf{A}^\top$.

Either way, one has $a_{ij}^\top = a_{ji}$ for all relevant pairs i, j.

Rows Before Columns

VERY Important Rule: Rows before columns!

This order really matters.

Reversing it gives a transposed matrix.

Exercise

Verify that the double transpose of any $m \times n$ matrix $\mathbf{A}$ satisfies $(\mathbf{A}^\top)^\top = \mathbf{A}$
— i.e., transposing a matrix twice recovers the original matrix.

Outline

Multiplying a Matrix by a Scalar

A scalar, usually denoted by a Greek letter,
is simply a real number $\alpha \in \mathbb{R}$.

The product of any $m \times n$ matrix $\mathbf{A} = (a_{ij})^{m \times n}$
and any scalar $\alpha \in \mathbb{R}$
is the new $m \times n$ matrix denoted by $\alpha\mathbf{A} = (\alpha a_{ij})^{m \times n}$,
each of whose elements αa_{ij} results
from multiplying the corresponding element a_{ij} of $\mathbf{A}$ by α.

Matrix Multiplication

The matrix product of two matrices $\mathbf{A}$ and $\mathbf{B}$
is defined (whenever possible) as the matrix $\mathbf{C} = \mathbf{AB} = (c_{ij})^{m \times n}$
whose element c_{ij} in row i and column j
is the inner product $c_{ij} = \mathbf{a}_i^\top \mathbf{b}_j$ of:

- the ith row vector $\mathbf{a}_i^\top$ of the first matrix $\mathbf{A}$;
- the jth column vector $\mathbf{b}_j$ of the second matrix $\mathbf{B}$.

Again: rows before columns!

Note that the resulting matrix product $\mathbf{C}$ must have:

- as many rows as the first matrix $\mathbf{A}$;
- as many columns as the second matrix $\mathbf{B}$.

Yet again: rows before columns!

Compatibility for Matrix Multiplication

Question: when is this definition of matrix product possible?

Answer: whenever $\mathbf{A}$ has as many columns as $\mathbf{B}$ has rows.

This condition ensures that every inner product $\mathbf{a}_i^\top \mathbf{b}_j$ is defined, which is true iff (if and only if) every row of $\mathbf{A}$ has exactly the same number of elements as every column of $\mathbf{B}$.

In this case, the two matrices $\mathbf{A}$ and $\mathbf{B}$ are compatible for multiplication.

Specifically, if $\mathbf{A}$ is $m \times \ell$ for some m, then $\mathbf{B}$ must be $\ell \times n$ for some n.

Then the product $\mathbf{C} = \mathbf{AB}$ is $m \times n$, with elements $c_{ij} = \mathbf{a}_i^\top \mathbf{b}_j = \sum_{k=1}^{\ell} a_{ik} b_{kj}$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Laws of Matrix Multiplication

Exercise

Verify that the following laws of matrix multiplication hold whenever the matrices are compatible for multiplication.

associative: $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$;

distributive: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ *and* $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$;

transpose: $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$.

shifting scalars: $\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B})$ *for all* $\alpha \in \mathbb{R}$.

Exercise

Let $\mathbf{X}$ *be any* $m \times n$ *matrix, and* $\mathbf{z}$ *any column n-vector.*

1. *Show that the matrix product* $\mathbf{z}^\top \mathbf{X}^\top \mathbf{X} \mathbf{z}$ *is well-defined, and that its value is a scalar.*
2. *By putting* $\mathbf{w} = \mathbf{Xz}$ *in the previous exercise regarding the value of the quadratic form* $\mathbf{w}^\top \mathbf{w}$, *what can you conclude about the value of the scalar* $\mathbf{z}^\top \mathbf{X}^\top \mathbf{X} \mathbf{z}$?

Exercise for Econometricians

Exercise

An econometrician has access to data series involving values

- y_t $(t = 1, 2, \ldots, T)$ *of one endogenous variable;*
- x_{ti} $(t = 1, 2, \ldots, T$ *and* $i = 1, 2, \ldots, k)$
  of k *different exogenous variables*
  — sometimes called explanatory variables or regressors.

The data is to be fitted into the linear regression model

$$y_t = \sum_{i=1}^{k} b_i x_{ti} + e_t$$

whose scalar constants b_i $(i = 1, 2, \ldots, k)$
are unknown regression coefficients,
and each scalar e_t *is the error term or residual.*

Exercise for Econometricians, Continued

1. Discuss how the regression model can be written in the form $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ for suitable column vectors **y, b, e**.
2. What are the dimensions of these vectors, and of the exogenous data matrix $\mathbf{X}$?
3. Why do you think econometricians use this matrix equation, rather than the alternative $\mathbf{y} = \mathbf{b}\mathbf{X} + \mathbf{e}$?
4. How can the equation accommodate the constant term α in the alternative equation $y_t = \alpha + \sum_{i=1}^{k} b_i x_{ti} + e_t$?

Matrix Multiplication Does Not Commute

The two matrices $\mathbf{A}$ and $\mathbf{B}$ commute just in case $\mathbf{AB} = \mathbf{BA}$.

Note that typical pairs of matrices DO NOT commute,
meaning that $\mathbf{AB} \neq \mathbf{BA}$ — i.e., the order of multiplication matters.

Indeed, suppose that $\mathbf{A}$ is $\ell \times m$ and $\mathbf{B}$ is $m \times n$,
as is needed for $\mathbf{AB}$ to be defined.

Then the reverse product $\mathbf{BA}$ is undefined
except in the special case when $n = \ell$.

Hence, for both $\mathbf{AB}$ and $\mathbf{BA}$ to be defined,
where $\mathbf{B}$ is $m \times n$, the matrix $\mathbf{A}$ must be $n \times m$.

But then $\mathbf{AB}$ is $n \times n$, whereas $\mathbf{BA}$ is $m \times m$.

Evidently $\mathbf{AB} \neq \mathbf{BA}$ unless $m = n$.

Thus all four matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{AB}$ and $\mathbf{BA}$ are $m \times m = n \times n$.

We must be in the special case
when all four are square matrices of the same dimension.

Matrix Multiplication Does Not Commute, II

Even if both $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices,
implying that both $\mathbf{AB}$ and $\mathbf{BA}$ are also $n \times n$,
one can still have $\mathbf{AB} \neq \mathbf{BA}$.

Here is a 2×2 example:

Example

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Exercise

For matrix multiplication,
why are there two different versions of the distributive law?

More Warnings Regarding Matrix Multiplication

Exercise

Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ *denote three matrices. Give examples showing that:*

1. *The matrix* $\mathbf{AB}$ *might be defined, even if* $\mathbf{BA}$ *is not.*
2. *One can have* $\mathbf{AB} = \mathbf{0}$ *even though* $\mathbf{A} \neq \mathbf{0}$ *and* $\mathbf{B} \neq \mathbf{0}$*.*
3. *If* $\mathbf{AB} = \mathbf{AC}$ *and* $\mathbf{A} \neq \mathbf{0}$*, it does not follow that* $\mathbf{B} = \mathbf{C}$*.*

Outline

Square Matrices

A square matrix has an equal number of rows and columns,
this number being called its dimension.

The (principal) diagonal
of a square matrix $\mathbf{A} = (a_{ij})_{n \times n}$ of dimension n
is the list $(a_{ii})_{i=1}^n = (a_{11}, a_{22}, \ldots, a_{nn})$ of its diagonal elements a_{ii}.

The other elements a_{ij} with $i \neq j$ are the off-diagonal elements.

A square matrix is often expressed in the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

with some extra dots along the diagonal.

Symmetric Matrices

A square matrix $\mathbf{A}$ is symmetric if it is equal to its transpose — i.e., if $\mathbf{A}^\top = \mathbf{A}$.

Example

The product of two symmetric matrices need not be symmetric.

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{but} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Two Exercises with Symmetric Matrices

Exercise

Let $\mathbf{x}$ be a column n-vector.

1. *Find the dimensions of $\mathbf{x}^\top \mathbf{x}$ and of $\mathbf{x}\mathbf{x}^\top$.*
2. *Show that one is a non-negative number which is positive unless $\mathbf{x} = \mathbf{0}$, and that the other is a symmetric matrix.*

Exercise

Let $\mathbf{A}$ be an $m \times n$-matrix.

1. *Find the dimensions of $\mathbf{A}^\top \mathbf{A}$ and of $\mathbf{A}\mathbf{A}^\top$.*
2. *Show that both $\mathbf{A}^\top \mathbf{A}$ and of $\mathbf{A}\mathbf{A}^\top$ are symmetric matrices.*
3. *Show that $m = n$ is a necessary condition for $\mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top$.*
4. *Show that $m = n$ with $\mathbf{A}$ symmetric is a sufficient condition for $\mathbf{A}^\top \mathbf{A} = \mathbf{A}\mathbf{A}^\top$.*

Diagonal Matrices

A square matrix $\mathbf{A} = (a_{ij})^{n \times n}$ is diagonal just in case
all of its off diagonal elements a_{ij} with $i \neq j$ are 0.

A diagonal matrix of dimension n can be written in the form

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{pmatrix} = \mathbf{diag}(d_1, d_2, d_3, \dots, d_n) = \mathbf{diag}\,\mathbf{d}$$

where the n-vector $\mathbf{d} = (d_1, d_2, d_3, \dots, d_n) = (d_i)_{i=1}^n$
consists of the diagonal elements of $\mathbf{D}$.

Obviously, any diagonal matrix is symmetric.

Multiplying by Diagonal Matrices

Example

Let $\mathbf{D}$ be a diagonal matrix of dimension n.

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ and $n \times m$ matrices, respectively.

Then $\mathbf{E} := \mathbf{AD}$ and $\mathbf{F} := \mathbf{DB}$ are well defined
as matrices of dimensions $m \times n$ and $n \times m$, respectively.

By the law of matrix multiplication, their elements are

$$e_{ij} = \sum\nolimits_{k=1}^{n} a_{ik} d_{kj} = a_{ij} d_{jj} \text{ and } f_{ij} = \sum\nolimits_{k=1}^{n} d_{ik} b_{kj} = d_{ii} b_{ij}$$

Thus, post-multiplying $\mathbf{A}$ by $\mathbf{D}$ is the column operation
of simultaneously multiplying every column $\mathbf{a}_j$ of $\mathbf{A}$
by its matching diagonal element d_{jj}.

Similarly, pre-multiplying $\mathbf{B}$ by $\mathbf{D}$ is the row operation
of simultaneously multiplying every row $\mathbf{b}_i^\top$ of $\mathbf{B}$
by its matching diagonal element d_{ii}.

Two Exercises with Diagonal Matrices

Exercise

Let $\mathbf{D}$ be a diagonal matrix of dimension n.
Give conditions that are both necessary and sufficient for each of the following:

1. $\mathbf{AD} = \mathbf{A}$ *for every* $m \times n$ *matrix* $\mathbf{A}$;
2. $\mathbf{DB} = \mathbf{B}$ *for every* $n \times m$ *matrix* $\mathbf{B}$.

Exercise

Let $\mathbf{D}$ be a diagonal matrix of dimension n, and $\mathbf{C}$ any $n \times n$ matrix.

An earlier example shows that one can have $\mathbf{CD} \neq \mathbf{DC}$ even if $n = 2$.

1. *Show that $\mathbf{C}$ being diagonal is a sufficient condition for $\mathbf{CD} = \mathbf{DC}$.*
2. *Is this condition necessary?*

Outline

The Identity Matrix

The identity matrix of dimension n is the diagonal matrix

$$\mathbf{I}_n = \mathbf{diag}(1, 1, \ldots, 1)$$

whose n diagonal elements are all equal to 1.

Equivalently, it is the $n \times n$-matrix $\mathbf{A} = (a_{ij})^{n \times n}$
whose elements are all given by $a_{ij} = \delta_{ij}$
for the Kronecker delta function $(i, j) \mapsto \delta_{ij}$
defined on $\{1, 2, \ldots, n\}^2$.

Exercise

Given any $m \times n$ matrix $\mathbf{A}$, verify that $\mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$.

Uniqueness of the Identity Matrix

Exercise

Suppose that the two $n \times n$ matrices $\mathbf{X}$ and $\mathbf{Y}$ respectively satisfy:

1. $\mathbf{AX} = \mathbf{A}$ *for every $m \times n$ matrix $\mathbf{A}$;*
2. $\mathbf{YB} = \mathbf{B}$ *for every $n \times m$ matrix $\mathbf{B}$.*

Prove that $\mathbf{X} = \mathbf{Y} = \mathbf{I}_n$.

(Hint: Consider each of the mn different cases where $\mathbf{A}$ (resp. $\mathbf{B}$) has exactly one non-zero element that is equal to 1.)

The results of the last two exercises together serve to prove:

Theorem

The identity matrix $\mathbf{I}_n$ is the unique $n \times n$-matrix such that:

- $\mathbf{I}_n\mathbf{B} = \mathbf{B}$ *for each $n \times m$ matrix $\mathbf{B}$;*
- $\mathbf{AI}_n = \mathbf{A}$ *for each $m \times n$ matrix $\mathbf{A}$.*

How the Identity Matrix Earns its Name

Remark

The identity matrix $\mathbf{I}_n$ *earns its name because it represents a* *multiplicative identity* *on the "algebra" of all* $n \times n$ *matrices.*

That is, $\mathbf{I}_n$ *is the unique* $n \times n$*-matrix with the property that* $\mathbf{I}_n\mathbf{A} = \mathbf{A}\mathbf{I}_n = \mathbf{A}$ *for every* $n \times n$*-matrix* $\mathbf{A}$*.*

Typical notation suppresses the subscript n in $\mathbf{I}_n$ that indicates the dimension of the identity matrix.

Outline

Permutations and Their Signs

Definition

Given $\mathbb{N}_n = \{1, \ldots, n\}$ where $n \geq 2$,
a permutation of $\mathbb{N}_n$ is a bijective mapping $\pi : \mathbb{N}_n \to \mathbb{N}_n$.
The family of all permutations Π includes:

- the identity mapping ι defined by $\iota(h) = h$ for all $h \in \mathbb{N}_n$;
- for each $\pi \in \Pi$, a unique inverse $\pi^{-1} \in \Pi$ for which $\pi^{-1} \circ \pi = \pi \circ \pi^{-1} = \iota$

Definition

1. Given any permutation π on $\mathbb{N}_n$, an inversion of π is a pair $(x, y) \in \mathbb{N}_n$ such that $i > j$ and $\pi(i) < \pi(j)$.
2. A permutation $\pi : \mathbb{N}_n \to \mathbb{N}_n$ is either even or odd according as it has an even or odd number of inversions.
3. The sign or signature of a permutation σ, denoted by $\text{sgn}(\pi)$, is defined as: $+1$ if π is even; and -1 if π is odd.

A Product Rule for the Signs of Permutations

Theorem

For any two permutations $\pi, \rho \in \Pi$, *one has*

$$\text{sgn}(\pi \circ \rho) = \text{sgn}(\pi)\,\text{sgn}(\rho)$$

The proof on the next slides uses the following:

Definition

First, define the signum or sign function

$$\mathbb{R} \setminus \{0\} \ni x \mapsto s(x) := \begin{cases} +1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Next, for each permutation $\pi \in \Pi$, let $S(\pi)$ denote the matrix whose elements satisfy $s_{ij}(\pi) = \begin{cases} -1 & \text{if } i > j \text{ and } \pi(i) < \pi(j) \\ +1 & \text{otherwise} \end{cases}$

Finally, let $\bigotimes S(\pi) := \prod_{i=1}^{n} \prod_{j=1}^{n} s_{ij}(\pi)$.

A Product Formula for the Sign of a Permutation

Lemma

For all $\pi \in \Pi$ *one has* $\text{sgn}(\pi) = \prod_{i>j} s(\pi(i) - \pi(j)) = \bigotimes S(\pi)$.

Proof.

Let $p := \#\{(i,j) \in \mathbb{N}_n \times \mathbb{N}_n \mid i > j \;\&\; \pi(i) < \pi(j)\}$ denote the number of inversions of π.

By definition, $\text{sgn}(\pi) = (-1)^p = \pm 1$ according as p is even or odd.

But the definitions on the previous slide imply that

$$\begin{aligned} p &= \#\{(i,j) \in \mathbb{N}_n \times \mathbb{N}_n \mid i > j \;\&\; s(\pi(i) - \pi(j)) = -1\} \\ &= \#\{(i,j) \in \mathbb{N}_n \times \mathbb{N}_n \mid s_{ij}(\pi) = -1\} \end{aligned}$$

Therefore

$$\begin{aligned} \text{sgn}(\pi) &= (-1)^p = \prod_{i>j} s(\pi(i) - \pi(j)) \\ &= \prod_{i=1}^{n} \prod_{j=1}^{n} s_{ij}(\pi) = \bigotimes S(\pi) \end{aligned}$$

□

Proving the Product Rule

Suppose the three permutations $\pi, \rho, \sigma \in \Pi$ satisfy $\sigma = \pi \circ \rho$.

Then $\dfrac{\text{sgn}(\sigma)}{\text{sgn}(\rho)} = \dfrac{\bigotimes S(\sigma)}{\bigotimes S(\rho)} = \prod_{i=1}^n \prod_{j=1}^n \dfrac{s_{ij}(\sigma)}{s_{ij}(\rho)}$

and also $s(\sigma(i) - \sigma(j)) = s(\pi(\rho(i)) - \pi(\rho(j)))$.

The definition of the two matrices $S(\sigma)$ and $S(\rho)$ implies that their elements satisfy $s_{ij}(\sigma)/s_{ij}(\rho) = 1$ whenever $i \leq j$.

In particular, $s_{ij}(\sigma)/s_{ij}(\rho) = 1$ unless both $i > j$ and also $s(\sigma(i) - \sigma(j)) = -s(\rho(i) - \rho(j))$.

Hence, given any $i, j \in \mathbb{N}_n$ with $i > j$, one has $s_{ij}(\sigma)/s_{ij}(\rho) = -1$ if and only if $s(\pi(\rho(i)) - \pi(\rho(j))) = -s(\rho(i) - \rho(j))$, or equivalently, if and only if:

either $\rho(i) > \rho(j)$ and $(\rho(i), \rho(j))$ is an inversion of π;

or $\rho(i) < \rho(j)$ and $(\rho(j), \rho(i))$ is an inversion of π.

Let p denote the number of inversions of the permutation π.

Then $\text{sgn}(\sigma)/\text{sgn}(\rho) = (-1)^p = \text{sgn}(\pi)$, implying that $\text{sgn}(\sigma) = \text{sgn}(\pi)\,\text{sgn}(\rho)$. □

Transpositions

For each disjoint pair $k, \ell \in \{1, 2, \ldots, n\}$,
the transposition mapping $i \mapsto \tau_{k\ell}(i)$ on $\{1, 2, \ldots, n\}$
is the permutation defined by

$$\tau_{k\ell}(i) := \begin{cases} \ell & \text{if } i = k; \\ k & \text{if } i = \ell; \\ i & \text{otherwise}; \end{cases}$$

Evidently $\tau_{k\ell} = \tau_{\ell k}$ and $\tau_{k\ell} \circ \tau_{\ell k} = \iota$, the identity permutation,
and so $\tau \circ \tau = \iota$ for every transposition τ.

It is also evident that τ has only one inversion, so $\text{sgn}(\tau) = -1$.

Exercise

Show that transpositions of more than two elements may not commute because, for example,

$$\tau_{12} \circ \tau_{23} = \pi^{231} \neq \tau_{23} \circ \tau_{12} = \pi^{312}$$

Permutations are Products of Transpositions

Theorem

Any permutation $\pi \in \Pi$ *on* $\mathbb{N}_n := \{1, 2, \ldots, n\}$
is the product of at most $n - 1$ *transpositions.*

We will prove the result by induction on n.

As the induction hypothesis,
suppose the result holds for permutations on $\mathbb{N}_{n-1}$.

Any permutation π on $\mathbb{N}_2 := \{1, 2\}$ is either the identity,
or the transposition τ_{12}, so the result holds for $n = 2$.

Proof of Induction Step

For general n, let:

- $j := \pi^{-1}(n)$ denote the element that π moves to the end;
- τ_{jn} denote the transposition that interchanges j and n.

By construction, the permutation $\pi \circ \tau_{jn}$
must satisfy $\pi \circ \tau_{jn}(n) = \pi(\tau_{jn}(n)) = \pi(j) = n$.

So the restriction $\tilde{\pi}$ of $\pi \circ \tau_{jn}$ to $\mathbb{N}_{n-1}$ is a permutation on $\mathbb{N}_{n-1}$.

By the induction hypothesis, for all $k \in \mathbb{N}_{n-1}$ one has

$$\tilde{\pi}(k) = \pi \circ \tau_{jn}(k) = \tau^1 \circ \tau^2 \circ \ldots \circ \tau^q(k)$$

where $q \leq n-2$ is the number of transpositions in the product.

For $p = 1, \ldots, q$, because τ^p interchanges only elements of $\mathbb{N}_{n-1}$,
one can extend its definition so that $\tau^p(n) = n$.

Then $\pi \circ \tau_{jn}(k) = \tau^1 \circ \tau^2 \circ \ldots \circ \tau^q(k)$ for $k = n$ as well, so

$$\pi = (\pi \circ \tau_{jn}) \circ \tau_{jn}^{-1} = \tau^1 \circ \tau^2 \circ \ldots \circ \tau^q \circ \tau_{jn}^{-1}$$

Hence π is the product of at most $q + 1 \leq n - 1$ transpositions.

This completes the proof by induction on n. □

Outline

Permutation Matrices: Definition

Definition

For each permutation $\pi \in \Pi$ on $\{1, 2, \ldots, n\}$, let $\mathbf{P}^\pi$ denote the unique associated n-dimensional permutation matrix which is derived by applying π to the rows of the identity matrix $\mathbf{I}_n$.

That is, for each $i = 1, 2, \ldots, n$,
the ith row vector of the identity matrix $\mathbf{I}_n$
is moved to become row $\pi(i)$ of $\mathbf{P}^\pi$.

This definition implies that the only nonzero element
in row i of $\mathbf{P}^\pi$ occurs not in column $j = i$,
as it would in the identity matrix,
but in column $j = \pi^{-1}(i)$ where $i = \pi(j)$.

Hence the matrix elements p^π_{ij} of $\mathbf{P}^\pi$
are given by $p^\pi_{ij} = \delta_{i,\pi(j)}$ for $i, j = 1, 2, \ldots, n$.

Permutation Matrices: Examples

Example

The two 2×2 permutation matrices are:

$$\mathbf{P}^{12} = \mathbf{I}_2; \quad \mathbf{P}^{21} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The $3! = 6$ permutation matrices in 3 dimensions are:

$$\mathbf{P}^{123} = \mathbf{I}_3; \quad \mathbf{P}^{132} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad \mathbf{P}^{213} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\mathbf{P}^{231} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad \mathbf{P}^{312} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}; \quad \mathbf{P}^{321} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Their signs are respectively $+1$, -1, -1, $+1$, $+1$ and -1.

Permutation Matrices: Exercise

Exercise

Suppose that π, ρ are permutations in Π, whose composition *is the function $\pi \circ \rho$ defined by*

$$\{1, 2, \ldots, n\} \ni i \mapsto (\pi \circ \rho)(i) = \pi(\rho(i)) \in \{1, 2, \ldots, n\}$$

Show that:

1. *the mapping $i \mapsto \pi(\rho(i))$ is a permutation on $\{1, 2, \ldots, n\}$;*
2. *the associated permutation matrices satisfy $\mathbf{P}^{\pi \circ \rho} = \mathbf{P}^{\pi} \mathbf{P}^{\rho}$.*

Transposition Matrices

A special case of a permutation matrix
is a transposition $\mathbf{T}_{h,i}$ of rows h and i.

As the matrix $\mathbf{I}$ with rows h and i transposed, it satisfies

$$(\mathbf{T}_{h,i})_{rs} = \begin{cases} \delta_{rs} & \text{if } r \notin \{h, i\} \\ \delta_{is} & \text{if } r = h \\ \delta_{hs} & \text{if } r = i \end{cases}$$

Exercise

Prove that:

1. *any transposition matrix* $\mathbf{T} = \mathbf{T}_{h,i}$ *is symmetric;*
2. $\mathbf{T}_{h,i} = \mathbf{T}_{i,h}$;
3. $\mathbf{T}_{h,i}\mathbf{T}_{i,h} = \mathbf{T}_{i,h}\mathbf{T}_{h,i} = \mathbf{I}$.

More on Permutation Matrices

Theorem

Any permutation matrix $\mathbf{P} = \mathbf{P}^\pi$ *satisfies:*

1. $\mathbf{P} = \prod_{s=1}^q \mathbf{T}^s$ *for the product of some collection of* $q \leq n-1$ *transposition matrices.*
2. $\mathbf{P}\mathbf{P}^\top = \mathbf{P}^\top\mathbf{P} = \mathbf{I}$

Proof.

The permutation π is the composition $\tau^1 \circ \tau^2 \circ \cdots \circ \tau^q$ of $q \leq n-1$ transpositions τ^s (for $s \in S := \{1, 2, \ldots, q\}$).

It follows that $\mathbf{P}^\pi = \prod_{s=1}^q \mathbf{T}^s$ where $\mathbf{T}^s = \mathbf{P}^{\tau^s}$ for $s \in S$.

Furthermore, because each $\mathbf{T}^s$ is symmetric, the transpose $(\mathbf{P}^\pi)^\top$ equals the reversed product $\mathbf{T}^q \cdots \mathbf{T}^2\mathbf{T}^1$.

But each transposition $\mathbf{T}^s$ also satisfies $\mathbf{T}^s\mathbf{T}^s = \mathbf{I}$, so

$$\mathbf{P}\mathbf{P}^\top = \mathbf{T}^1\mathbf{T}^2 \cdots \mathbf{T}^q\mathbf{T}^q \cdots \mathbf{T}^2\mathbf{T}^1 = \mathbf{T}^1\mathbf{T}^2 \cdots \mathbf{T}^{q-1}\mathbf{T}^{q-1} \cdots \mathbf{T}^2\mathbf{T}^1 = \mathbf{I}$$

by induction on q, and similarly $\mathbf{P}^\top\mathbf{P} = \mathbf{I}$. □

Outline

A First Elementary Row Operation

Suppose that row r of the $m \times m$ identity matrix $\mathbf{I}_m$
is multiplied by a scalar $\alpha \in \mathbb{R}$, leaving all other rows unchanged.

The result is the $m \times m$ diagonal matrix $\mathbf{S}_r(\alpha)$
whose diagonal elements are all 1,
except the (r, r) element which is α.

Hence the elements of $\mathbf{S}_r(\alpha)$ satisfy

$$(\mathbf{S}_r(\alpha))_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq r \\ \alpha\delta_{ij} & \text{if } i = r \end{cases}$$

Exercise

*For the particular $m \times m$ matrix $\mathbf{S}_r(\alpha)$
and the general $m \times n$ matrix $\mathbf{A}$,
show that the transformed $m \times n$ matrix $\mathbf{S}_r(\alpha)\mathbf{A}$
is the result of multiplying row r of $\mathbf{A}$ by the scalar α,
leaving all other rows unchanged.*

A Second Elementary Row Operation

Suppose a multiple of α times row q of the identity matrix $\mathbf{I}_m$ is added to its rth row, leaving all the other $m-1$ rows unchanged.

The resulting $m \times m$ matrix $\mathbf{E}_{r+\alpha q}$ equals $\mathbf{I}_m$, but with an extra non-zero element equal to α in the (r, q) position.

Its elements therefore satisfy

$$(\mathbf{E}_{r+\alpha q})_{ij} = \begin{cases} \delta_{ij} & \text{if } i \neq r \\ \delta_{rj} + \alpha \delta_{qj} & \text{if } i = r \end{cases}$$

Exercise

For the particular $m \times m$ matrix $\mathbf{E}_{r+\alpha q}$ and the general $m \times n$ matrix $\mathbf{A}$, show that the transformed $m \times n$ matrix $\mathbf{E}_{r+\alpha q}\mathbf{A}$ is the result of adding the multiple of α times its row q to the rth row of matrix $\mathbf{A}$, leaving all other rows unchanged.

Levi-Civita Symbols

For any $n \in \mathbb{N}$, define the set $\mathbb{N}_n := \{1, 2, \ldots, n\}$ of the first n natural numbers.

Definition

The Levi-Civita symbol $\epsilon_{\mathbf{j}} = \epsilon_{j_1 j_2 \ldots j_n} \in \{-1, 0, 1\}$ is defined for all (ordered) lists $\mathbf{j} = (j_1, j_2, \ldots, j_n) \in (\mathbb{N}_n)^n$.

Its value depends on whether the mapping $\mathbb{N}_n \ni i \mapsto j_i \in \mathbb{N}_n$ is an even or an odd permutation of the ordered list $(1, 2, \ldots, n)$, or is not a permutation at all. Specifically,

$$\epsilon_{\mathbf{j}} = \epsilon_{j_1 j_2 \ldots j_n} := \begin{cases} +1 & \text{if } i \mapsto j_i \text{ is an even permutation} \\ -1 & \text{if } i \mapsto j_i \text{ is an odd permutation} \\ 0 & \text{if } i \mapsto j_i \text{ is not a permutation} \end{cases}$$

The Levi-Civita Matrix

Definition

Given the Levi-Civita mapping $\mathbb{N}_n \ni i \mapsto j_i \in \mathbb{N}_n := \{1, 2, \ldots, n\}$, the associated $n \times n$ Levi-Civita matrix $\mathbf{L_j}$ has elements defined by

$$(\mathbf{L_j})_{rs} = (\mathbf{L}_{j_1 j_2 \ldots j_n})_{rs} := \delta_{j_r, s}$$

This implies that the rth row of $\mathbf{L_j}$ equals row j_r of the matrix $\mathbf{I}_n$.

That is, $\mathbf{L_j} = (\mathbf{e}_{j_1}, \mathbf{e}_{j_2}, \ldots, \mathbf{e}_{j_n})^\top$ is the $n \times n$ matrix produced by stacking the n row vectors $\mathbf{e}_{j_r}^\top$ $(r = 1, 2, \ldots, n)$ of the canonical basis on top of each other, with repetitions allowed.

For a general $n \times n$ matrix $\mathbf{A}$, the matrix $\mathbf{L_j A} = \mathbf{L}_{j_1 j_2 \ldots j_n} \mathbf{A}$ is the result of stacking the n row vectors $\mathbf{a}_{j_r}^\top$ $(r = 1, 2, \ldots, n)$ of $\mathbf{A}$ on top of each other, with repetitions allowed.

Specifically, $(\mathbf{L_j A})_{rs} = a_{j_r s}$ for all $r, s \in \{1, 2, \ldots, n\}$.